

A P P L I C A T I O N

for

UNITED STATES LETTERS PATENT

on

Site-Specific Recombination in Eukaryotes
and Constructs Useful Therefor

by

Stephen O'Gorman and
Geoffrey Wahl

Sheets of Drawings: Two

Docket No.: SALK 2190
Salk File No.: S97025

Attorneys

Gray Cary Ware & Freidenrich
4365 Executive Drive, Suite 1600
San Diego, California 92121-2189

03919501-032897
248220-19561580

Site-Specific Recombination in Eukaryotes
and Constructs Useful Therefor

FIELD OF THE INVENTION

5 The present invention relates to methods for manipulating chromosomal sequences in cells by site-specific recombination promoted by recombinases. In a particular aspect, the present invention relates to methods for producing embryonic stem cells bearing nucleic acid sequences that have been rearranged by a site-specific recombinase expressed from a construct controlled by a tissue-specific promoter (e.g., a germline specific
10 promoter). In another aspect, the present invention relates to methods for producing embryonic stem cells bearing nucleic acid sequences that have been rearranged by a site-specific recombinase expressed from a construct controlled by a conditional promoter.

15 BACKGROUND OF THE INVENTION

The analysis of gene function has increasingly come to require the production of subtle, tissue-specific, and conditional mutations in animals and plants. Although there are a number of methods for engineering subtle
20 mutations in embryonic stem (ES) cells (Hasty et al. (1991) *Nature* 350:243-246, Askew et al. (1993) *Mol Cell Biol* 13:4115-4124), the use of site-specific recombinases to remove the selectable marker that permits isolation of homologously recombined ES cell clones has become
25 increasingly prevalent (Kitamoto et al. (1996) *Biochem Biophys Res Commun* 222:742-747, Fiering et al. (1993) *Proc Natl Acad Sci U S A* 90:8469-8473, Schwenk et al. (1995) *Nucleic Acids Res* 23:5080-5081; Gu et al. (1993) *Cell* 73:1155-1164; Sailer et al. (1996) *Taniguchi Symposia on*

08919501.082897

Ln
a1

Brain Sciences, eds. Nakanishi et al. (Japan Scientific Press), pp. 89-98).

Site-specific recombinases represent the best method for creating tissue-specific and conditional mutations in animals and plants, being employed first to remove the selectable marker to create a functionally wild-type allele, and then to inactivate the allele mosaically in animals and plants by removing some essential component in a tissue-specific or conditional manner (Gu et al. (1994) *Science* **265**:103-106; Kuhn et al. (1995) *Science* **269**:1427-1429). Current protocols for using excissive site-specific recombination to remove selectable markers include transiently transfecting ES cell clones with a recombinase expression vector (Gu et al. (1993) *Cell* **73**:1155-1164), microinjecting fertilized oocytes containing the recombinant allele with a recombinase expression vector (Kitamoto et al. (1996) *Biochem Biophys Res Commun* **222**:742-747; Araki et al. (1995) *Proc Natl Acad Sci U S A* **92**:160-164), or breeding animals and plants containing the recombinant allele to animals and plants, respectively, containing a recombinase transgene (Schwenk et al. (1995) *Nucleic Acids Res* **23**:5080-5081; Lewandoski et al. (1997) *Curr Biol* **7**:148-151). Each of these approaches requires an investment of some combination of time, resources, and expertise over that required to generate animals and plants with homologously recombined alleles. The most commonly employed method, the secondary transfection of homologously recombined ES cell clones with a recombinase expression vector, additionally requires extended culture time that may decrease their potential to enter the germline.

In principle, marker excision would be substantially simplified through the use of ES cells containing recombinase nucleic acid constructs that were expressed in the germline, but not to an appreciable extent

08919501.082697

in the ES cells themselves or somatic tissues of animals and plants. The lack of ES cell expression would mean that targeting vectors containing selectable markers flanked by recombinase target sites could be used to isolate homologous recombinants without fear that the marker would be excised during culture. Robust recombinase expression in gametes would mean that the marker would be excised in at least some of the progeny of ES cell chimeras. Only a single step would be required to isolate subtle mutations and, if two different recombinase systems were employed, conditional and tissue-specific alleles could be produced with similar improvements in efficiency. A germline-specific recombinase nucleic acid construct could also be used to deliver recombined target nucleic acid constructs to the early embryo (Lewandoski et al. (1997) *Curr Biol* 7:148-151), so long as the recombined target was not detrimental to the terminal stages of spermatogenesis.

Previous reports have shown that expression of nucleic acid constructs containing the proximal promoter of the mouse protamine 1 (mP1) locus is restricted to haploid spermatids in mature mice (Peschon et al. (1987) *Proc Natl Acad Sci U S A* 84:5316-5319; Behringer et al. (1988) *Proc Natl Acad Sci U S A* 85:2648-2652), although low levels of ectopic expression may occur in some mature tissues (Behringer et al. (1988) *Proc Natl Acad Sci USA* 85:2648-2652). Inclusion of the mP1 promoter does not guarantee expression in the male germline, however, for although nucleic acid constructs containing the mP1 promoter and the SV40 T-antigen coding sequence were transcribed, the message was not translated at detectable levels in spermatids (Behringer et al. (1988) *Proc Natl Acad Sci USA* 85:2648-2652).

Accordingly, there is a need in the art for methods to modulate expression of recombined target nucleic

2025 RELEASE UNDER E.O. 14176

acid sequences in the early embryo. In addition, there is a need in the art for tissue-specific and conditional recombinatory tools to create transgenic animals and plants. These and other needs in the art are addressed by
5 the present invention.

BRIEF DESCRIPTION OF THE INVENTION

The present invention meets the need in the art for modulating expression of recombined target nucleic acid sequences to the early embryo. The present invention
10 further meets the need in the art for tissue-specific and conditional recombinatory tools to create transgenic animals and plants. Thus, in accordance with the present invention, it has been discovered that nucleic acid constructs encoding a germline specific promoter
15 operatively associated with a recombinase coding sequence lead to efficient recombination of a target nucleic acid construct in the male germline, but not in other tissues. This suggests that such nucleic acid constructs could be used for the efficient production of embryos bearing
20 conditional, genetically lethal alleles. It has additionally been discovered that ES cell lines generated from one of these transgenic lines could be used in combination with targeting vectors that contained loxP-flanked selectable markers to isolate homologous
25 recombinants containing the marker and functional loxP sites.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 illustrates a schematic of P2Bc and P2Br alleles. The positions of the PCR primers used (5'P and
30 3'P) are indicated on the diagrams of the P2Bc and P2Br alleles.

00019501.08239
00019501.08239

Figure 2 depicts the targeting of the *hoxb-1* locus in ProCre ES cells using a targeting vector that contains a loxP-flanked selectable marker. Top, schematic of the wild-type *hoxb-1* locus showing the positions of the two exons (open boxes), the position of a 5' NruI site and flanking BamHI restriction endonuclease sites, and PCR primers (triangles) that amplify a 204 bp product from the wild-type allele that includes the NruI site. Middle, the predicted organization of homologously recombined *hoxb-1* allele in which a neomycin cassette (NEO), flanked by loxP sites (L), has been inserted into the NruI site shown in the top diagram. The insertion creates a novel BamHI site and the same PCR primers now amplify a 1600 bp product. Bottom: the predicted structure of the recombined allele shown in the middle panel after Cre-mediated excision of the neomycin cassette to leave a single loxP site in place of the NruI site of the wild-type allele. Amplification with the same primers now yields a 268 bp product.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, there are provided nucleic acid constructs comprising a germline-specific promoter operatively associated with a recombinase coding sequence.

As used herein, the term "promoter" refers to a specific nucleotide sequence recognized by RNA polymerase, the enzyme that initiates RNA synthesis. The promoter sequence is the site at which transcription can be specifically initiated under proper conditions. The recombinase nucleic acid(s), operatively linked to the suitable promoter, is(are) introduced into the cells of a suitable host, wherein expression of the recombinase nucleic acid(s) is(are) controlled by the promoter.

Germline-specific promoters contemplated for use in the practice of the present invention include the protamine 1 gene promoter, the protamine 2 gene promoter, the spermatid-specific promoter from the c-kit gene
5 (Albanesi et al. (1996) *Development* **122**(4):1291-1302), the sperm-specific promoter from angiotensin-converting enzyme (Howard et al. (1993) *Mol Cell Biol* **13**(1):18-27; Zhou et al. (1995) *Dev Genet* **16**(2):201-209), oocyte specific promoter from the ZP1 gene, oocyte specific promoter from
10 the ZP2 gene, oocyte specific promoter from the ZP3 gene (Schickler et al. (1992) *Mol Cell Biol* **12**(1):120-127), and the like.

In addition to the above-described germline-specific promoters, tissue-specific promoters specific to
15 plants are also contemplated for use in the practice of the present invention, including, for example, the LAT52 gene promoter from tomato, the LAT56 gene promoter from tomato, the LAT59 gene promoter from tomato Eyal et al. (1995) *Plant Cell* **7**(3):373-384), the pollen-specific promoter of
20 the Brassica S locus glycoprotein gene (Dzelzkalns et al. (1993) *Plant Cell* **5**(8):855-863), the pollen-specific promoter of the NTP303 gene (Weterings et al. (1995) *Plant J* **8**(1):55-63), and the like.

Recombinases contemplated for use in the practice
25 of the present invention include Cre recombinase, FLP recombinase, the R gene product of *Zygosaccharomyces* (Onouchi et al. (1995) *Mol Gen Genet* **247**(6):653-660), and the like.

Presently preferred constructs contemplated for
30 use in the practice of the present invention include ProCre (comprising the protamine 1 gene promoter operatively associated with Cre recombinase), ProFLP (comprising the protamine 1 gene promoter operatively associated with FLP

08919501.082897

recombinase), ProR (comprising the protamine 1 gene promoter operatively associated with the R gene product of *Zygosaccharomyces*), and the like.

5 In accordance with another embodiment of the present invention, there are provided nucleic acid constructs comprising a conditional promoter or a tissue-specific promoter operatively associated with a recombinase coding sequence.

Promoters contemplated for control of expression
10 of recombinase nucleic acid(s) employed in accordance with this aspect of the present invention include inducible (e.g., minimal CMV promoter, minimal TK promoter, modified MMLV LTR), constitutive (e.g., chicken β -actin promoter, MMLV LTR (non-modified), DHFR), and/or tissue specific
15 promoters.

Conditional promoters contemplated for use in the practice of the present invention comprise transcription regulatory regions that function maximally to promote transcription of mRNA under inducing conditions. Examples
20 of suitable inducible promoters include DNA sequences corresponding to: the *E. coli* lac operator responsive to IPTG (see Nakamura et al., *Cell*, 18:1109-1117, 1979); the metallothionein promoter metal-regulatory-elements responsive to heavy-metal (e.g., zinc) induction (see Evans
25 et al., U.S. Patent No. 4,870,009), the phage T7lac promoter responsive to IPTG (see Studier et al., *Meth. Enzymol.*, 185: 60-89, 1990; and U.S. #4,952,496), the heat-shock promoter; the TK minimal promoter; the CMV minimal promoter; a synthetic promoter; and the like.

30 Exemplary constitutive promoters contemplated for use in the practice of the present invention include the CMV promoter, the SV40 promoter, the DHFR promoter, the

202501-08239

mouse mammary tumor virus (MMTV) steroid-inducible promoter, Moloney murine leukemia virus (MMLV) promoter, elongation factor 1 α (EF1 α) promoter, albumin promoter, APO A1 promoter, cyclic AMP dependent kinase II (CaMKII) promoter, keratin promoter, CD3 promoter, immunoglobulin light or heavy chain promoters, neurofilament promoter, neuron specific enolase promoter, L7 promoter, CD2 promoter, myosin light chain kinase promoter, HOX gene promoter, thymidine kinase (TK) promoter, RNA Pol II promoter, MYOD promoter, MYF5 promoter, phosphoglycerokinase (PGK) promoter, Stf1 promoter, Low Density Lipoprotein (LDL) promoter, chicken β -actin promoter (used in conjunction with ecdysone response element) and the like.

As readily understood by those of skill in the art, the term "tissue specific" refers to the substantially exclusive initiation of transcription in the tissue from which a particular promoter, which drives expression of a given gene, is derived (e.g., expressed only in T-cells, endothelial cells, smooth muscle cells, and the like). Exemplary tissue specific promoters contemplated for use in the practice of the present invention include the GH promoter, the NSE promoter, the GFAP promoter, neurotransmitter promoters (e.g., tyrosine hydroxylase, TH, choline acetyltransferase, ChAT, and the like), promoters for neurotropic factors (e.g., a nerve growth factor promoter, NT-3, BDNF promoters, and the like), and so on.

In accordance with yet another embodiment of the present invention, there are provided embryonic stem cells containing a nucleic acid construct as described herein.

As readily understood by those of skill in the art, the above-described constructs can be introduced into a variety of animal species, such as, for example, mouse, rat, rabbits, swine, ruminants (sheep, goats and cattle), humans, poultry, fish, and the like. Transgenic

amphibians, insects, nematodes, and the like, are also contemplated. Members of the plant kingdom, such as, for example, transgenic mono- and dicotyledonous species, including important crop plants, i.e., wheat, rice, maize, soybean, potato, cotton, alfalfa, and the like, are also contemplated.

For example, pluripotential ES cells can be derived from early preimplantation embryos, preferably the ova are harvested between the eight-cell and blastocyst stages. ES cells are maintained in culture long enough to permit integration of the promoter-recombinase nucleic acid construct(s). The cells are then either injected into a host blastocyst, i.e., the blastocoel of the host blastocyst, or cocultured with eight-cell to morula-stage ova, i.e, zona-free morula, so that transfected ES cells are preferentially incorporated into the inner cell mass of the developing embryo. With blastocyst injection, transgenic offspring are termed "chimeric," as some of their cells are derived from the host blastocyst and some transfected ES cells. The host embryos are transferred into intermediate hosts or surrogate females for continuous development.

The transformation procedure for plants usually relies on the transfer of a transgene carrying a particular promoter construct via the soil bacterium *Agrobacterium tumefaciens*. Transformation vectors for this procedure are derived from the T-DNA of *A. tumefaciens*, and transgenes are stably incorporated into the nuclear genome. The activity of the transgenes can then be monitored in the regenerated plants under different conditions. In this way, many promoter elements that are involved in complex regulatory pathways such as light responsiveness or tissue specificity have been defined.

00019501.082897

Alternatively, direct (i.e., vectorless) gene transfer systems are also contemplated including chemical methods, electroporation, microinjection, biolistics, and the like. Protoplasts isolated from the plants can be
5 obtained by treatment with cell wall degrading enzymes. DNA can be introduced into plant protoplasts by a number of physical techniques including electroporation and polyethylene glycol treatment in the presence of $MgCl_2$. The method of choice for rapid promoter analyses in plants
10 is the biolistic method. This technique involves the delivery of the particular DNA construct into plant cells by microprojectiles, i.e., nucleic acid(s) coated or precipitated by tungsten or gold. This method is not limited to any particular plant species or tissue type.
15 Preferably, this method would allow quantitative analysis of transformation if appropriate selectable markers are included.

In a preferred embodiment, the genome of embryonic stem cells according to the invention comprise a
20 transcriptionally active selectable marker flanked by two recombination target sites. It is especially preferred that the recombinase encoded by the recombinase coding sequence operatively associated with a germline-specific promoter is selective for the recombination target sites
25 flanking said selectable marker.

Optionally, embryonic stem cells according to the invention may further comprise one or more of:

a nucleic acid fragment flanked by two recombination target sites, wherein said recombination
30 target sites are different than the recombination target sites which flank said selectable marker,

a nucleic acid construct comprising a conditional promoter operatively associated with a recombinase coding sequence,

5

10

25

30

30 passaging the genome derived from said embryonic
stem cells through gametogenesis (i.e., spermatogenesis or
oogenesis).

Excission of marker as contemplated herein can cause a variety of end results, e.g., deletion of the marker or a nucleic acid sequence, gain of function or loss of function, replacement of function, and the like, as well as modulation of any one or more of these results.

Functions which are contemplated to be manipulated include regulating body size and growth rate, including recombining gene constructs which contain various growth hormone gene sequences. Other productivity traits that are targets include altering the properties or proportions of caseins, lactose, or butterfat in milk, increased resistance to viral and bacterial diseases (i.e., "constitutive immunity" or germ-line transmission of specific, recombined antibody genes), more efficient wool production, and the like. Other functions which are contemplated to be modulated include development of lines of transgenic animals and plants for use in directing expression of transgenes encoding biologically active human proteins.

Agronomic traits which are contemplated to be modulated by use of the present invention include tolerance to biotic and abiotic stresses, increased resistance to herbicides, pest damage, and viral, bacterial, and fungal diseases, improvement of crop quality (i.e., increase in nutritional value of food and feed), reduction of post-harvest losses, improvement of suitability and enlargement of the spectrum for processing (i.e., altered quantity and composition of endogenous properties, production of new compounds of plant or non-plant origin such as biopolymers or pharmaceutical substances).

In accordance with a still further embodiment of the present invention, there are provided methods for the production of recombinant alleles, said method comprising:

00019501-002897

passaging the genome derived from said embryonic
5 stem cells through gametogenesis.

As readily recognized by those of skill in the art, nucleic acid fragments can be introduced into ES cells by a variety of techniques, e.g., by homologous recombination, random insertion, retroviral insertion, site specific-mediated recombination, and the like.

Nucleic acid fragments contemplated for use herein include fragments containing an essential portion of a gene of interest.

In accordance with yet another embodiment of the present invention, there are provided methods for the production of recombinant alleles, said method comprising:

introducing at least one recombinase responsive construct
into embryonic stem cells as described herein,
wherein said construct(s) comprise(s) a
20 nucleic acid fragment and a selectable marker,
wherein said selectable marker is flanked by
a first pair of recombination target sites, and
wherein said nucleic acid fragment is
flanked by a second pair of recombination target
25 sites,

passaging the genome derived from said embryonic stem cells through gametogenesis.

In a presently preferred aspect, the first pair of recombination target sites is recognized by a recombina-
30 se which is expressed under the control of a germline-specific promoter and said second pair of

recombination target sites is recognized by a recombinase which is expressed under the control of a conditional promoter or a tissue specific promoter.

Optionally, the embryonic stem cells employed
5 herein can further comprise a second nucleic acid construct selected from constructs comprising a conditional promoter operatively associated with a recombinase coding sequence, a construct comprising a tissue-specific promoter operatively associated with a recombinase coding sequence,
10 and the like.

In accordance with still another embodiment of the present invention, there are provided methods for the conditional assembly of functional gene(s) for expression in eukaryotic cells by recombination of individual inactive
15 gene segments from one or more gene(s) of interest,

wherein each of said segments contains at least one recombination target site, and

wherein at least one of said segments contains at least two recombination target sites,

20 said method comprising:

introducing said individual inactive gene segments into an embryonic stem cell as described herein, thereby providing a DNA which encodes a functional gene of interest, the expression
25 product of which is biologically active, upon passage of the genome derived from said stem cells through gametogenesis.

For assembly of functional genes from inactive gene segments, see, for example, US Patent No. 5,654,182,
30 incorporated herein by reference in its entirety.

In accordance with a still further embodiment of the present invention, there are provided methods for the

00919501.082897

combining embryonic stem cells that include
nucleic acid construct according to the invention with host
5 pluripotent ES cells derived from early preimplantation
embryos, and

10 In accordance with yet another embodiment of the
present invention, there are provided methods for the
generation of recombinant plants, said method comprising
transforming plant zygotes with nucleic acid constructs
according to the invention and allowing the zygote to
15 develop.

The objective of the current work with ProCre nucleic acid constructs was to determine the potential of germline-specific promoters to implement efficient approaches utilizing site-specific recombinases to generate an array of sophisticated mutations in mammals and plants. The data shows that it is possible to create recombinase nucleic acid constructs that are expressed at high levels in the germ line but not to a functionally significant extent in either ES cells or embryonic or adult somatic tissues. Homologous recombinants with a selectable marker can be isolated in ES cells that contain promoter-recombinase nucleic acid constructs. Transgenic animals and plants bearing the promoter-recombinase nucleic acid constructs and a target allele transmit the recombined target to their progeny at high frequencies. These results establish the principle that mammals and plants containing loci that have been homologously recombined and then subsequently site-specifically recombined can be generated simply by using ES cells with a suitable recombinase nucleic acid constructs for the initial targeting. By this mechanism, alleles containing a single recombinase target

site and a mutation of interest can be produced in the progeny of ES cell chimeras without any investment of time, expertise, or resources over that required to create an allele that still contains a selectable marker. The
 5 paradigm has obvious utility in the production of subtle and conditional mutations that require generation of alleles with minimal structural alterations. Because the presence and transcriptional activity of selectable markers can contribute to phenotypes in an unanticipated and
 10 unwanted manner (Fiering et al. (1995) *Genes Dev* 9:2203-2213); Olson et al. (1996) *Cell* 85:1-4), the approach will also be useful for generating null alleles.

Expression of the endogenous mPl locus (Hecht et al. (1986) *Exp Cell Res* 164:183-190), and mPl-driven
 15 nucleic acid constructs (Behringer et al. (1988) *Proc Natl Acad Sci U S A* 85:2648-2652; Braun et al. (1989) *Nature* 337:373-376; Zambrowicz et al. (1993) *Proc Natl Acad Sci U S A* 90:5071-5075) is restricted to haploid spermatids. Expression of mPl nucleic acid construct expression
 20 typically begins at haploid stages, and both RNA (Caldwell and Handel (1991) *Proc Natl Acad Sci U S A* 88:2407-2411) and proteins (Braun et al. (1989) *Nature* 337:373-376) diffuse through the spermatogenic syncytium. The result is a highly efficient recombination of target alleles and the
 25 segregation of recombinase and target nucleic acid constructs in the first generation.

Cre-mediated recombination proved to be highly testis-specific in ProCre mice. It is clear that the nucleic acid constructs are not expressed in the inner cell
 30 mass or in other early embryonic tissues. Cells from preimplantation embryos intermingle extensively and the embryo as a whole is derived from a small number of cells (Beddington et al. (1989) *Development* 106:37-46; Soriano and Jaenisch (1986) *Cell* 46:19-29). If ProCre nucleic acid

00919501-002007
 268280-1051620

constructs recombined target sequences during preimplantation stages, at least a few percent of the cells in many tissues would contain the P2Br allele and Southern and PCR analyses showed that this was not the case. The
5 ectopic Cre activity seen in some ProCre strains probably resulted from low levels of recombinase expression in later embryos or mature tissues, a finding consistent with the expression patterns of other mPl-driven nucleic acid constructs. Northern analyses have failed to reveal the
10 expression of mPl-containing nucleic acid constructs in a variety of mature tissues (Peschon et al. (1987) *Proc Natl Acad Sci USA* **84**:5316-5319; Behringer et al. (1988) *Proc Natl Acad Sci U S A* **85**:2648-2652; Peschon et al. (1989) *Ann N Y Acad Sci* **564**:186-197; Zambrowicz et al. (1993) *Proc*
15 *Natl Acad Sci U S A* **90**:5071-5075), but nucleic acid constructs containing the mPl promoter and the SV40 T-antigen led to the consistent development of tumors of the petrosal bone and right cardiac atrium (Behringer et al. (1988) *Proc Natl Acad Sci U S A* **85**:2648-2652).

20 PCR assays represent a very sensitive assay for whether sufficient levels of Cre protein were produced to effect recombination. Importantly, they measured the cumulative level of recombination, for events that occurred at any stage of development are likely to have been
25 propagated to, and might be amplified in, descendant populations. The highest level of ectopic recombination was that observed in cardiac ventricular tissue of strain which generated a signal approximately equivalent to that expected if the ratio between recombined and
30 unrecombined alleles were 1:104. The activities observed in other strains were considerably lower than this, and one strain did not show any ectopic activity. None of the strains showed evidence of recombination in the cardiac atria and the petrosal bone was not examined. These assays
35 did not rule out the possibility that higher levels of recombination occur in tissues that were not examined or

00049501.002097

These low levels of ectopic activity suggest that
5 mpl-driven recombinase nucleic acid constructs could be
used for the production of embryos containing genetically
lethal alleles. Some alleles created by homologous
recombination in ES cells will prove to be lethal in
heterozygotes, as was the case for an mRNA editing mutation
10 of the GluR2 glutamate receptor subunit (Brusa et al.
(1995) *Science* **270**:1677-1680). Germline transmission would
be restricted to rare chimeras in which the level of
chimerism was low enough in tissues affected by the
mutation to allow survival and high enough in the germline
15 to allow transmission. This problem could be circumvented
by creating recombinase-conditional mutations in ES cells
bearing mpl-recombinase nucleic acid constructs, or by
making the same mutations in standard ES cells and then
introducing the mpl-recombinase nucleic acid construct by
20 breeding. So long as the recombined version of the allele
did not adversely impact terminal stages of
spermatogenesis, embryos containing the recombined allele
could be efficiently produced. Embryos containing
recombined nucleic acid constructs can also be produced
25 through the activity of Cre nucleic acid constructs that
are expressed during early embryogenesis from the human
cytomegalovirus minimal promoter (Schwenk et al. (1995)
Nucleic Acids Res **23**:5080-5081), the adenovirus EIIa
promoter (Lakso et al. (1992) *Proc Natl Acad Sci U S A*
30 **89**:6232-6236), or the zP3 promoter (Lewandoski et al.
(1997) *Curr Biol* **7**:148-151). ProCre and zP3 nucleic acid
constructs have the advantage of delivering a recombined
allele to the zygote, guaranteeing that all cells in the
derived embryos will contain the allele.

ProCre ES cells are but one of many different kinds of recombinase-bearing ES cells that could significantly shorten the time and effort required for a wide variety of genetic manipulations in mice. The most obvious of these are complementary ProFLP ES cells in which the FLP recombinase was derived from *S. cerevisiae* (Broach and Hicks (1980) *Cell* 21:501-508) or another species (Kuhn et al. (1995) *Science* 269:1427-1429). Conceptually distinct from these but perhaps as generically useful would be ES cells bearing inducible recombinase nucleic acid constructs that would facilitate temporal control of recombinase expression in ES cells, chimeras, and their progeny to generate site-specifically recombined alleles (Araki et al. (1992) *J Mol Biol* 225:25-37; No et al. (1996) *Proc Natl Acad Sci U S A* 93:3346-3351; Logie and Stewart (1995) *Proc Natl Acad Sci U S A* 92:5940-5944; Feil et al. (1996) *Proc Natl Acad Sci U S A* 93:10887-10890). Finally, fusion genes that led to recombinase expression in specific tissues could be used to address specific research objectives.

The invention will now be described in greater detail by reference to the following non-limiting examples.

Example 1

Mammalian DNA Constructs

A 652 bp fragment of the mP1 promoter (SEQ ID NO:1; Peschon et al. (1989) *Annals of the New York Academy of Sciences* 186-197) was isolated by PCR using genomic DNA from CCE ES cells (Robertson et al. (1986) *Nature* 323:445-448) as a template. This fragment was fused to a Cre coding sequence modified to contain a consensus translation start site (Kozak (1986) *Cell* 44:283-292), 11 codons for a human c-myc epitope (Evan et al. (1985) *Mol Cell Biol* 5:3610-3616), 7 codons for a minimal SV40 nuclear

08919501 082397

localization signal (Kalderon et al. (1984) *Cell* 39:499-509) and the polyadenylation signal from pIC-Cre in the plasmid pOG304M (SEQ ID NO:4). The Cre expression plasmid pOG231 was prepared by fusing a Cre coding sequence
5 modified from pIC-Cre (Gu et al. (1993) *Cell* 73:1155-1164), and containing the same translation start and nuclear localization signal, to the synthetic intron and CMV promoter of pOG44 (O'Gorman et al. (1991) *Science* 251:1351-1355).

10 A plasmid, pOG277 (SEQ ID NO:6), containing a loxP-flanked neomycin cassette was prepared by inserting a wild-type loxP site (SEQ ID NO:5; Hoess et al. (1982) *Proc Natl Acad Sci U S A* 79:3398-402) into pBSKS (Stratagene) and then cloning the neomycin expression cassette from
15 pMCIneo-polyA (Thomas et al. (1987) *Cell* 51:503-512) between interations of this loxP site. The *hoxb-1* targeting construct consisted of the PGK-TK cassette from pPNT (Tybulewicz et al. (1991) *Cell* 65:1153-63), and 1.4kb and 10.2kb of sequences 5' and 3' to an *Nru I* site 800 bp
20 5' to the *hoxb-1* transcriptional start site isolated from a 129 strain genomic library (Stratagene). The loxP-flanked neo cassette from pOG277 was inserted into the *NruI* site. The pOG277 neomycin cassette and a β -GAL sequence was inserted into the first exon of the large
25 subunit of RNA polymerase II (RP2) (Ahearn et al. (1987) *J. Biol. Chem.* 262:10695-10705) to create the P2Bc allele (Figure 1). Cre-mediated recombination of the P2Bc allele results in the deletion of the neomycin cassette (Neo) of P2Bc that is flanked by two loxP sites, leaving a single
30 loxP site and fusing the B-Gal coding sequence to the initial codons of the RNA polymerase II coding sequence. Recombination increases the size of a *Pst I* fragment recognized by the RP2 probe, which is external to the targeting vector used, indicated by the shaded box below
35 each allele.

00019501.082897

Example 2Production of transgenic mice

Fertilized oocytes obtained from matings of 129/SvJae (Simpson et al. (1997) *Nat Genet* **16**:19-27) and
5 BALB/c X C57BL/6 F1 mice were used for pronuclear
injections of the Protamine-Cre fusion gene from pOG304M
according to standard protocols (Hogan et al. *Manipulating
the Mouse Embryo: The Manual*, Coldspring Harbor Press
(1994), pg. 497). Production of ES cells and homologous
10 recombinants: Heterozygous ProCre 129/SvJae males were
mated to 129/SvEms-+^{Ter?}/J females (Simpson et al. (1997) *Nat
Genet* **16**:19-27) to produce blastocysts that were cultured
according to standard protocols (Robertson (1987)
Teratocarcinomas and embryonic stem cells, a practical
15 approach, eds. E. J. Robertson (IRL Press), pp. 71-112).
The sex (King et al. (1994) *Genomics* **24**:159-68) and ProCre
status of each line were determined by PCR assays.
Molecular analyses: Tail biopsy genomic DNA was used for
hybridization assays or PCR assays to identify ProCre and
20 P2Bc/r mice. PCR reactions for the detection of ectopic
Cre activity used 100 ng of genomic DNA as a template to
amplify a P2Br-specific product using a 5' primer from the
RP2 promoter and a 3' primer from the β -GAL coding sequence
(Figure 1). Thirty cycles of amplification were done in a
25 total volume of 100 μ l using 300 ng of each primer, 3 mM
MgCl₂, 1.5 units of Taq polymerase, and an annealing
temperature of 60°C. Southern blots of reaction products
were hybridized with a probe specific for the P2Br reaction
product (SEQ ID Nos:2 and 3).

00919501 082897

Example 3ProCre Nucleic Acid Constructs Efficiently Recombine
Target Alleles

A total of nine founder animals with ProCre
5 nucleic acid constructs were obtained from injections of a
Protamine-Cre fusion gene. Two lines were derived from
injections of 129SvJae (Simpson et al. (1997) *Nat Genet*
16:19-27) embryos, and seven from injections of CB6F2
embryos. The 129/SvJae lines and three randomly selected
10 hybrid lines were examined in detail. To determine whether
ProCre nucleic acid constructs would efficiently recombine
a target allele, males were generated that contained a
ProCre nucleic acid construct and a target for Cre-mediated
recombination. This "P2Bc" (Pol II, β -GAL, conditional)
15 target (Figure 1) was created using homologous
recombination in ES cells to insert a loxP-flanked neomycin
cassette and a β -GAL coding sequence into the first exon of
the locus coding for the large subunit of RNA
polymerase II. Cre-mediated recombination of the loxP
20 sites was expected to delete the intercalated sequences,
creating "P2Br" allele (Pol II, β -Gal, recombined).

These males were mated to wild-type females and
the resulting progeny were examined by Southern blotting to
determine if they inherited the P2Bc or the P2Br allele,
25 and to additionally determine the segregation pattern of
ProCre nucleic acid constructs and P2Br alleles. Southern
blot of Pst I digested tail biopsy DNA's from a +/P2Bc,
+/ProCre male (sire) and four of his progeny by a wild-type
female probed with n RP2 probe (top) and then reprobed with
30 a Cre probe (bottom). The large majority of transmitted
target alleles were Cre-recombined P2Br alleles (Table 1).
ProCre nucleic acid constructs and recombined target
alleles segregated independently in the first generation;
approximately 50% of mice that inherited a P2Br allele also
35 inherited their male parent's ProCre nucleic acid

00919501 082897

construct. All RP2 mutant alleles in the progeny were P2Br, and some progeny inherit a P2Br allele without inheriting ProCre nucleic acid construct. Mouse 4 did not contain a ProCre nucleic acid construct and is homozygous wild-type at the RP2 locus. These data establish that ProCre nucleic acid constructs efficiently recombine the P2Bc allele in the male germline and that the recombined P2Br alleles and ProCre nucleic acid constructs segregate in the first generation. Because significantly more than 25% of the progeny inherited recombined target alleles, recombination either occurred during diploid stages of spermatogenesis or Cre generated during haploid stages of spermatogenesis was distributed among spermatids through cytoplasmic bridges (Braun et al. (1989) *Nature* 337:373-376), effecting recombination in spermatids that did not themselves contain a ProCre nucleic acid construct.

The progeny of matings between ProCre males and +/-P2Bc females were also examined to determine if male gametes from ProCre mice delivered enough Cre to zygotes to effect Cre-mediated recombination of a target sequence. Of 96 progeny examined by Southern blotting, none contained a Cre-recombined P2Br allele.

It has also been discovered that a loxP-flanked neo cassette in the glutamate receptor R6 subunit locus is efficiently recombined by ProCre nucleic acid constructs in mice.

Example 4

ProCre Nucleic acid construct Expression is Highly Tissue-Specific

Genomic DNAs from ten different tissues of five- to seven-week old males that contained both a ProCre nucleic acid construct and a P2Bc target allele were analyzed in Southern blots. Southern blots were prepared

of Pst I digested DNA from testes (T) and one other tissue (K, kidney; B, brain; S, spleen) of males heterozygous for one of four ProCre nucleic acid constructs and the P2Bc allele. Testis DNA from each male shows a P2Br allele signal, in addition to those generated by the wild-type RP2 (WT) and P2Bc alleles. Other tissues show only the WT and P2Bc signals. Only the testis samples showed signal indicating Cre-mediated recombination of the target. The intensity of the P2Br signal relative to that of the wild-type allele ranged from 10% to 22% for different ProCre strains and did not correlate with the ProCre nucleic acid construct copy number. The copy number of ProCre nucleic acid constructs varied among lines showing similar levels of recombination in testis. For example, restriction patterns and densitometric analyses showed that line 58 contained a single copy of the ProCre nucleic acid construct, yet showed virtually the same testis recombination signal as line containing more than 100 copies. This variability is similar to results obtained with other mP1 promoter-driven nucleic acid constructs (Peschon et al. (1987) *Proc Natl Acad Sci USA* 84:5316-5319; Zambrowicz et al. (1993) *Proc Natl Acad Sci USA* 90:5071-5075).

As a more sensitive measure of ectopic recombination, PCR amplifications were performed on the same samples. The amplification primers were expected to produce a 325 bp product from the recombined target and a 1.4 kb fragment from the unrecombined allele (Figure 1). The assay was expected to measure the cumulative level of recombination, for any P2Br alleles formed during transient expression of Cre during development would be preserved and perhaps amplified in descendant cells. Low levels of ectopic recombination product were observed in some tissues of all ProCre lines except for one. A southern blot of PCR amplification products of the P2Br allele utilized tissues from a male heterozygous for the ProCre nucleic acid

08919501-082897

construct and the P2Bc allele. DNA from 10 different tissues was amplified using primers and conditions that produced a 350 bp product from the recombined, P2Br allele. Each lane contains 10% of the reactions, except for the
5 testis reactions, which were diluted 500 (T5), 250 (T2), and 100 (T1) fold prior to loading, and a liver reconstruction control (C), which was diluted 1:100 before loading. The highest level of ectopic activity was observed in cardiac ventricular muscle of mice; in these
10 samples the ectopic signal was more than 100 fold lower than that observed in testis. Three strains showed much lower levels of recombination in brain tissue, and strain 75 additionally showed ectopic activity in spleen. Despite the difficulty of quantifying PCR results, these
15 data clearly indicate that ectopic activity occurred at very low levels in most tissues of most ProCre lines.

Example 5

Isolation of Homologously Recombined ProCre ES Cell Clones Using Targeting Vectors with a loxP-Flanked 20 Selectable Marker

Four male +/-ProCre ES cell lines were established from 129/Sv strain ProCre transgenic mice. In preliminary experiments, passage 5 cells from one of these lines (PC3) were used to generate three male chimeras with between 50
25 and 95% coat color chimerism. In matings with C57BL/6 females, two of these male chimeras have sired a total of 11 pups, all bearing the Agouti coat color signifying germline transmission of the ES cell genome, and 6 of 9 pups genotyped additionally contained the line 70 ProCre
30 nucleic acid construct. The frequency of germline transmission has not yet been determined, nor has it been determined whether competency for germline transmission will persist in homologously recombined ProCre ES cells at later passages.

08919501.032897
463280.10561680

To determine if homologously recombined ProCre ES cell clones could be isolated using targeting vectors that contained a loxP-flanked selectable marker, two transfections were done using variants of a targeting
5 vector in which a loxP-flanked neomycin cassette was inserted into an Nru I site in the *hoxb-1* locus promoter (Figure 2). A Southern blot of BamHI-digested genomic DNAs were harvested from a 96-well plate from 10 doubly-selected ES cell clones and hybridized with a probe (shown in Figure
10 2) which is external to the targeting construct. All samples show the 7.5 kb band from the wild-type allele and four clones additionally show the 6 kb band predicted to result from homologous recombination. In these transfections, 12 of 62 (19%) PC3- and 10 of 56 (18%)
15 PCs-derived clones that were ganciclovir and G418-resistant (Mansour et al. (1988) *Nature* 336:348-352) were found to be homologously recombined. In two parallel transfections of CCE cells (Robertson et al. (1986) *Nature* 323:) with the same vectors, 32 of 93 (34%) and 15 of 132 (11%) clones
20 were homologously recombined. The total numbers of G418-resistant clones recovered from ProCre ES cell transfections were reduced relative to the parallel CCE transfections. This may be attributable to both Cre-mediated excision of the neomycin cassette and to the
25 fact that the transfections were done under electroporation conditions optimized for CCE cells.

Because it was formally possible that the homologously recombined clones contained inactive loxP sites, five homologously recombined PC3 ES cell clones and
30 the parental PC3 cell line using the primers shown in Figure 2 were either mock transfected or transiently transfected with the pOG231 Cre expression vector. For the transient transfection assay, DNA was harvested 48 hours after transfection and used in PCR assays to assess whether
35 the loxP sites in the recombinant clones could be recombined by Cre. In all cases a clear recombination

signal was observed in the pOG231 transfected sample. The recombinant clones and parental cell lines show the 204 bp amplification product of the wild-type allele, and the recombinant clones additionally show a 1600 bp product (1600) resulting from amplification across the neomycin cassette and a nonspecific 1100 bp amplification product (NS). The pOG231-transfected recombinant clones show an additional 268 bp product signaling the Cre-mediated excision of the neomycin cassette from the recombinant alleles of some cells. Experiments were also done to assess the stability of the loxP-flanked neo cassette in ProCre ES cells. Five recombinant clones were grown in the presence of G418 for two weeks, and then aliquots of each were grown either in the presence or absence of G418 for a further 10 days. PCR assays were performed to determine if Cre-recombined alleles were present in any of these samples and none was observed in the mock transfected controls. These data suggest that there is not enough Cre activity to significantly influence either the ability to isolate recombinant clones or the stability of the selectable markers in those clones, establishing that the loxP sites in these clones were functional.

To determine if there was any detectable Cre activity in ProCre ES cells, aliquots of two lines (PC3 and PC5) were transiently transfected with the targeting vector used to create the P2Bc allele. DNA was recovered 48 hours after transfection and used for PCR amplifications of the P2Br plasmid molecules that would be generated by extrachromosomal Cre-mediated recombination. Small amounts of recombination product were seen in both ProCre ES cell transfections, and none was observed in parallel samples of CCE ES cells. This shows that the ProCre ES cell lines express sufficient Cre to recombine some extrachromosomal targets when the latter are present at high copy numbers.

Example 6
Plant DNA Constructs

To define sequences in the LAT52 and LAT59 promoters involved in expression in pollen, proximal
5 promoters were constructed employing a series of linker substitution mutants using the particle bombardment system (Klein et al. (1987) *Nature* 327:70-73; Twell et al. (1989b) *Plant Physiol* 91:1270-1274). These experiments were performed by cobombarding the test plasmids (luciferase
10 [LUC]- recombinase fusions) with reference plasmids (β -glucuronidase [GUS] fusions). The latter served as a control for bombardment variability and allowed comparisons to be made between independent bombardments.

The context of the -100 promoter in LAT52 and the
15 -115 promoter in LAT59 was chosen because these promoters appeared to be the minimal regions that still conferred high levels (25% relative to the available full-length promoter) of pollen-specific expression (Twell et al. (1991) *Gen Dev* 5:496-507). These minimal promoters were
20 then fused to the Cre coding sequence operatively linked to the *luc* gene (Ow et al. (1986) *Science* 234:856-858) coding region, and the resulting plasmids served as a basis for creating the nucleic acid constructs. The LAT52 linker substitutions were performed in p52LUC, which contain
25 entire LAT52 5' untranslated region (5' UTR). A series of six 9- to 10-bp-long linker substitutions were made in p52LUC, spanning the region -84 to -29 (52LS1 to 52LS6).

Example 7
Tissue Specificity in Plants

30 The results obtained by transient expression in pollen and in transgenic plants provided information on the effect of the various constructs on expression in pollen but not on their effect on tissue specificity. A tobacco

cell culture, TXD (maintained as described by Howard et al. (1992) *Cell* 68:109-118), was, therefore, added as an additional component of the transient assay system. The TXD cell culture was initiated from tobacco mesophyll cells and therefore represents somatic tissue, as opposed to the gametophytic tissue represented by pollen. Cells in culture were chosen, rather than intact tissue, as the somatic tissue source because such cells superficially resemble pollen in that they can be spread out as a monolayer on a plate before bombardment.

In this experiment, translation fusions between the *luc* coding region and either the CaMV 35S promoter drove strong expression in cell culture but negligible expression in pollen, whereas the LAT52 promoter showed the opposite pattern of strong activity in pollen and negligible activity in cell culture. Thus, the transient assay system mimics the expression pattern observed for these promoters in transgenic plants (Twell et al. (1991) *Genes Dev* 5:496-507). This differential expression provided us with a tool with which to address tissue specificity.

Example 8

Plant Transformation and Analysis of Transgenic Plants

Constructs cloned into pBin19 were introduced into tomato (*Lycopersicon esculentum* cv VF36) by *Agrobacterium tumefaciens* LBA4404 as previously described (McCormick (1991b) Transformation of tomato with *Agrobacterium tumefaciens*, In Plant Tissue Culture Manual, K. Linsey, Ed B6:1-9). At least 20 independent transformants were obtained for each construct.

For β -glucuronidase (GUS) assays, 5 to 20 μ L of pollen, pooled from several flowers of the same plant, was ground directly in Eppendorf tubes in 50 to 100 μ L of GUS

extraction buffer (Jefferson et al. (1987) *EMBO* 6:3901-3907) using a Teflon-tipped homogenizer driven by a drill. Expression in pollen was measured by fluorometrically assaying GUS activity in supernatants of pollen extracts
 5 using 2mM 4-methylumbelliferyl β -D-glucuronide (Sigma) as substrate (Jefferson et al. (1987) *EMBO* 6:3901-3907). GUS activity was corrected for variation in total protein content using a bicinchoninic acid protein assay kit (Pierce, Rockford, IL).

10 Expression in leaves, flowers, stems, roots, and seed was tested histochemically by staining with 5-bromo-4-chloro-3-indolyl β -D-glucuronide (Molecular Probes, Eugene, OR) as described previously (Jefferson et al. (1987) *EMBO* 6:3901-3907). Expression in leaves was also analyzed
 15 fluorometrically as given previously.

Example 9

Transient Transformation of Tobacco Pollen and Cell Culture

Pollen spread out as a monolayer was bombarded
 20 essentially as previously described (Twell et al. (1991) *Genes Dev* 5:496-507), except that gold was substituted for tungsten and only 1 μ g of test plasmid and used per plate. TXD cell culture (maintained as described by Howard et al. (1992) *Cell* 68:109-118) was spread out similarly as a
 25 monolayer (1 mL of a 50-mL stationary culture per plate) and bombarded as previously described. Between six and 12 independent bombardments were performed for each construct. In each experiment, the test plasmid was cobombarded with a reference plasmid: pB1223 (Clontech, Palo Alto, CA) was
 30 used for assays of all constructs in tobacco cell culture; pLAT59-12 (Twell et al. (1990) *Development* 109:705-713) for assays of LAT52 and LAT56 constructs in tobacco pollen; pLAT56-12 (Twell et al. (1990) *Development* 109:705-713) for

0091501 "002897"

assays of LAT59 constructs in tobacco pollen. Processing of the tissue after ~ 15 to 17 hr and analysis of GUS and LUC activity were as described previously (Twell et al. (1991) *Genes Dev* 5:496-507). Transient expression was
5 reported as "relative LUC activity," which represents the ratio between the test (LUC) and the reference (GUS) plasmids.

While the invention has been described in detail with reference to certain preferred embodiments thereof, it
10 will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

08919501-082297